JOURNAL OF BACTERIOLOGY, June 1991, p. 3446–3455
0021-9193/91/113446-10$02.00/0

Vol. 173, No. 11

Evolutionary Relationships among Eubacteria, Cyanobacteria, and Chloroplasts: Evidence from the *rpoC1* Gene of *Anabaena* sp. Strain PCC 7120

KRISTIN J. BERGSLAND* AND ROBERT HASELKORN

Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

Received 4 February 1991/Accepted 29 March 1991

RNA polymerases of cyanobacteria contain a novel core subunit, γ, which is absent from the RNA polymerases of other eubacteria. The genes encoding the three largest subunits of RNA polymerase, including γ, have been isolated from the cyanobacterium *Anabaena* sp. strain PCC 7120. The genes are linked in the order *rpoB*, *rpoC1*, *rpoC2* and encode the β, γ, and β′ subunits, respectively. These genes are analogous to the *rpoBC* operon of *Escherichia coli*, but the functions of *rpoC* have been split in *Anabaena* between two genes, *rpoC1* and *rpoC2*. The DNA sequence of the *rpoC1* gene was determined and shows that the γ subunit corresponds to the amino-terminal half of the *E. coli* β′ subunit. The γ protein contains several conserved domains found in the largest subunits of all bacterial and eukaryotic RNA polymerases, including a potential zinc finger motif. The spliced *rpoC1* gene from spinach chloroplast DNA was expressed in *E. coli* and shown to encode a protein immunologically related to *Anabaena* γ. The similarities in the RNA polymerase gene products and gene organizations between cyanobacteria and chloroplasts support the cyanobacterial origin of chloroplasts and a divergent evolutionary pathway among eubacteria.

RNA polymerase in prokaryotes generally consists of a catalytic core of four subunits ($\beta\beta'\alpha_2$) and a dissociable sigma factor, which confers promoter specificity (7, 16, 44). The basic structure of the enzyme was believed to be the same in all eubacteria until the RNA polymerase of the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120 was purified (37). The cyanobacterial core RNA polymerase was found to contain, in addition to a β, a β′, and two α's, a novel core subunit of 70 kDa designated γ, which is absent from the RNA polymerases of other eubacteria. Western immunoblotting with antiserum to γ has shown that a serologically related γ protein is present in over 30 different cyanobacteria, representing the five major taxonomic subgroups (4a, 36). Thus, the γ subunit is a common feature of cyanobacterial RNA polymerases.

Anti-γ serum cross-reacts with the *Escherichia coli* β′ subunit protein, suggesting homology between these subunits (36). A region of DNA homologous to part of the *E. coli* gene encoding β′ (*rpoC*) has been isolated from another cyanobacterium, *Nostoc commune* UTEX 584 (47). Sequencing of this region showed that homology to *E. coli rpoC* is split between two linked genes, *rpoC1* and *rpoC2*, in the *Nostoc* DNA (47). The DNAs of several plant chloroplasts have likewise been found to contain regions of sequence homology with *E. coli rpoC* (19, 20, 30, 38, 39). In chloroplasts, the blocks of homology are also distributed between two genes, *rpoC1* and *rpoC2*. These genes are linked and are thought to encode the β′ and β″ subunits of chloroplast RNA polymerase, respectively.

The exact subunit composition of chloroplast RNA polymerase is unclear. Spinach chloroplast RNA polymerase was found to contain seven prominent polypeptides, some of which were serologically related to subunits of the *E. coli* RNA polymerase (26). DNA sequences potentially encoding protein equivalents to the *E. coli* α and β subunits (*rpoA* and *rpoB*) have been found in chloroplast genomes, in addition to the *rpoC* homologs (19, 20, 30, 39, 40). In a few cases, these genes have been shown to be expressed in chloroplasts. The protein product of the *rpoA* gene has been identified in maize and pea chloroplasts (33, 34). Polypeptides corresponding to the *rpoB* and *rpoC2* genes have also been found in purified preparations of maize chloroplast RNA polymerase (18). The presence of RNA transcripts from the *rpoBC1C2* genes in spinach chloroplasts has been demonstrated (19), but the protein products of these genes have not yet been identified. In this report, we present evidence that the spinach chloroplast *rpoC1* gene encodes a protein immunologically related to the *Anabaena* γ subunit.

In this study, we have isolated the complete *rpoBC1C2* region from *Anabaena* sp. strain PCC 7120 and shown that it encodes the β, γ, and β′ subunits of RNA polymerase. We have determined the DNA sequence of the *rpoC1* gene, encoding γ, in order to examine the function of γ and its relationship to other RNA polymerase subunit proteins. We compare the sequence of the *Anabaena* γ subunit with the β′ subunit of chloroplast and other bacterial RNA polymerases and discuss how these results provide evidence for the divergence of cyanobacteria and chloroplasts from the eubacterial lineage.

MATERIALS AND METHODS

Bacterial strains, plasmids, and phages. Strains, plasmids, and phages are listed in Table 1 along with the respective sources or references.

Isolation of *Anabaena* sp. strain PCC 7120 RNA polymerase genes. The 2.8- and 2.3-kb *Eco*RI fragments containing parts of the *E. coli rpoB* and *rpoC* genes, respectively, were isolated from pGB218 (2), which was a gift from C. Squires. The fragments were labeled with [α-^{32}P]dCTP (3,000 Ci/mmol; New England Nuclear) by the random priming method (11) and used to probe a Southern blot of *Anabaena* sp. strain PCC 7120 chromosomal DNA digested with vari-

* Corresponding author.

TABLE 1. Bacterial strains, plasmids, and phages

Strain or plasmid	Relevant characteristics	Source or reference
E. coli		
MC1061	Propagation of plasmids	6
BL21(DE3)pLysS	IPTG-inducible expression of cloned genes directed by T7 RNA polymerase	42
TG1	Propagation of M13 phage	Amersham
DH5α	Host for pUC plasmids	Bethesda Research Laboratories
Plasmids and phages		
pGB218	pBR322 derivative, contains *E. coli rpoBC* genes	2
pBR328	Cloning vector; Amp^r Tet^r Cam^r	Boehringer Mannheim Biochemicals
pKH100	pBR328 + 1.3-kb *Hin*dIII-*Eco*RI fragment of *rpoB* from *Anabaena* sp. strain PCC 7120	This work
pWB79	Cosmid vector used for *Anabaena* DNA libraries; Amp^r	5
pID9	Cosmid pWB79 with 35-kb insert including the *rpoBC1C2* region from *Anabaena* sp. strain PCC 7120	This work
pT7tet18 and -19	Vectors for expression of genes by T7 RNA polymerase; contain a T7 promoter upstream of multiple cloning site; Tet^r	4
pT181	pT7tet18 + 11.4-kb *Xba*I fragment containing the *Anabaena rpoBC1C2* region; forward orientation	This work
pT182	pT7tet18 + 11.4-kb *Xba*I fragment; reverse orientation	This work
pT195	pT7tet19 + 2.8-kb *Hin*dIII-*Kpn*I fragment containing the *Anabaena rpoC1* gene; forward orientation	This work
pT185	pT7tet18 + 2.8-kb *Hin*dIII-*Kpn*I fragment; reverse orientation	This work
pUC18 and -19	Vector for subcloning and sequencing; Amp^r	48
pSocS4	Spinach chloroplast *Sal*I fragment 4 in pUC8; contains *rpoBC1C2* region	19
pUC1928	pUC19 + 2.8-kb *Hin*dIII-*Sca*I fragment from pSocS4; spinach chloroplast *rpoC1* gene	This work
pUC1921	pUC19 + 2.1-kb *Hin*dIII-*Kpn*I fragment from mp1921; chloroplast *rpoC1* gene with intron deleted	This work
M13mp19	Vector used for oligonucleotide-directed mutagenesis	48
mp1928	M13mp19 + 2.8-kb *Hin*dIII-*Sca*I fragment with *rpoC1* gene from spinach chloroplast	This work
mp1921	mp1928 with intron deleted	This work
pATH3	Vector for making *trpE* gene fusions; Amp^r	23
pATH301	pATH3 + 2.1-kb *Xmn*I-*Kpn*I fragment; *trpE* fusion with *Anabaena rpoC1*	This work
pATH302	pATH3 + 2.1-kb *Cla*I-*Kpn*I fragment from pUC1921; *trpE* fusion with intron-deleted chloroplast *rpoC1* gene	This work
pATH303	pATH3 + 2.8-kb *Cla*I-*Kpn*I fragment from pUC1928; *trpE* fusion with chloroplast *rpoC1* containing intron	This work

ous restriction enzymes. Enzymes were purchased from Boehringer Mannheim Biochemicals and used according to the manufacturer's instructions. Agarose gel electrophoresis and Southern blotting onto GeneScreen Plus membrane (DuPont-NEN Research Products) were done by standard methods (27). Hybridizations and washes were carried out according to the GeneScreen Plus product guide. Heterologous hybridizations and washes were done at 60°C.

The 1.3-kb *Eco*RI-*Hin*dIII fragment containing part of the *rpoB* gene was isolated from *Anabaena* sp. strain PCC 7120 chromosomal DNA by making a two-step size-directed DNA library. A 100-μg sample of DNA was first digested with *Eco*RI and separated on an agarose gel. Fragments of 5.8 to 6.8 kb were isolated and digested with *Hin*dIII, and fragments of 1.0 to 1.5 kb were recovered. This DNA was ligated with pBR328, which had been digested with *Eco*RI and *Hin*dIII, and then used to transform *E. coli* MC1061. pKH100 was identified by screening the resultant plasmid miniprep DNAs by Southern blotting and hybridization with the *E. coli* 2.8-kb *rpoB* fragment. The 1.3-kb insert of pKH100 was used to probe two libraries of *Anabaena* sp. strain PCC 7120 chromosomal DNA in the cosmid vector pWB79 (5). The libraries were made from partial digests of the chromosomal DNA with either *Cpf*I or *Hin*dIII and contained inserts in the range of 31 to 45 kb (5). Homologous hybridizations and washes were done at 65°C.

Subcloning and expression of *rpo* genes in *E. coli*. Smaller fragments were subcloned from the cosmid pID9 into pT7tet18 or pT7tet19 (4) for expression of proteins in *E. coli*. After restriction enzyme digestion of pID9, fragments were prepared by separation on an agarose gel followed by electrophoresis onto an NA-45 DEAE membrane (Schleicher & Schuell) and elution at 68°C with high-salt buffer (1 M NaCl, 20 mM Tris [pH 8], 0.1 mM EDTA). The 11.4-kb *Xba*I fragment was ligated into the *Xba*I site of pT7tet18 to make pT181 and pT182, which have the fragment in opposite orientations. The 2.8-kb *Hin*dIII-*Kpn*I fragment was cloned into the *Hin*dIII and *Kpn*I sites of pT7tet18 and pT7tet19 to make pT185 and pT195, which have the fragment in opposite orientations with respect to the T7 promoter. The plasmids were used to transform *E. coli* BL21(DE3)pLysS (42), and the strains were grown on LB medium (27) containing 10 μg each of chloramphenicol and tetracycline per ml. For preparation of proteins, the strains were grown overnight in 5 ml of LB containing chloramphenicol and tetracycline and then diluted 1:100 into 30 ml of fresh medium. The cultures were grown at 37°C with shaking to an optical density at 600 nm of 0.4; 1 ml was removed, and isopropyl-β-D-thiogalactopyran-

oside (IPTG; Sigma Chemical Co.) was added to a final concentration of 0.5 mM. Following a further 60-min incubation at 37°C with shaking, the bacteria were collected by centrifugation and resuspended in 100 μl of distilled H_2O. Then 100 μl of 2× final sample buffer (25) was added, and the samples were boiled for 3 min. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis was performed as described by Laemmli (25), with a 10% separating gel and a 4.5% stacking gel; 50 μl of each sample was loaded per gel.

Western blot analysis was carried out essentially as described by Tortorello and Dunny (43). Electrophoretic transfer of proteins to nitrocellulose was done in 20 mM Tris base–150 mM glycine–20% methanol at 70 V overnight. Antibodies to *Anabaena* sp. strain PCC 7120 core RNA polymerase and the γ subunit were previously described (36). Blots were developed by using goat anti-rabbit immunoglobulin G-horseradish peroxidase conjugate (Cappel) and 4-chloro-1-naphthol (Sigma Chemical Co.).

DNA sequence analysis of *rpoC1*. The DNA sequence of both strands of the 2.8-kb *Hin*dIII-*Kpn*I fragment in pT185 was determined. DNA sequencing was done by the method of Sanger et al. (35), using plasmid DNA templates and a Sequenase kit (United States Biochemicals Corp.). [α-^{35}S] dATP (1,350 Ci/mmol) was purchased from New England Nuclear. Templates were prepared from miniprep plasmid DNA according to Kraft et al. (24). Nested deletions for sequencing were generated from pT185, previously digested with either *Hin*dIII or *Kpn*I, using BAL 31 exonuclease (New England BioLabs). The method was essentially that of Heinrich et al. (15) except that the deleted fragments were isolated from an agarose gel by electrophoresis onto an NA-45 membrane (see above) and then recloned into pUC18 or pUC19 for sequencing with the universal M13/pUC 17-mer forward sequencing primer. DNA sequence data were analyzed by using the DNAlysis program written by William Buikema for the Apple Macintosh and the sequence software package from the University of Wisconsin Genetics Computer Group.

Expression of the spinach chloroplast *rpoC1* gene. The 2.8-kb *Hin*dIII-*Sca*I fragment containing the spinach chloroplast *rpoC1* gene was isolated from pSocS4 (19) and ligated into the *Hin*dIII and *Sma*I sites of pUC19 to make pUC1928. pSocS4 DNA was a gift from G. Hudson. The 2.8-kb *Hin*dIII-*Kpn*I fragment from pUC1928 was isolated and ligated into the *Hin*dIII and *Kpn*I sites of M13mp19 to make mp1928. The 756-bp intron of the chloroplast *rpoC1* gene was precisely deleted from mp1928 by using an oligonucleotide-directed in vitro mutagenesis kit purchased from Amersham. The oligonucleotide used to create the deletion was a 32-mer, with a 16-base arm on either side of the sequence to be deleted. This oligomer had the sequence 5′-GCCTAG CAAATGAAAAATCGCAGTATACTAGG-3′ and was complementary to nucleotides 5412 to 5427 and 6184 to 6199 of the DNA sequence of the spinach chloroplast *rpoBC1C2* region reported by Hudson et al. (19). The precision of the deletion in mp1921 was verified by sequence analysis using a 17-base primer (5′-AGGATTGGATTTCATAT-3′) complementary to nucleotides 6246 to 6262 of the sequence of Hudson et al. (19).

The 2.1-kb *Hin*dIII-*Kpn*I fragment from mp1921 was isolated from a miniprep of the phage replicative-form DNA and cloned into the *Hin*dIII and *Kpn*I sites of pUC19 to make pUC1921. To put the chloroplast *rpoC1* gene in the same translational reading frame as the *trpE* gene in pATH3 (23), the *rpoC1* genes in pUC1921 and pUC1928 were isolated on *Cla*I-*Kpn*I fragments and cloned into the *Bam*HI site of pATH3. The *Cla*I end was filled in by using Klenow polymerase (Pharmacia), and the *Kpn*I end was made blunt by using T4 DNA polymerase (Boehringer Mannheim Biochemicals). The *Bam*HI ends of the vector were likewise filled in by Klenow polymerase and then treated with alkaline phosphatase (Boehringer Mannheim Biochemicals) to prevent religation. Following transformation into *E. coli* MC1061, the orientations of the inserts were verified by restriction digestion. pATH302 and pATH303 contain the 2.1- and 2.8-kb *Cla*I-*Kpn*I fragments from pUC1921 and pUC1928, respectively.

A *trpE* fusion with the *Anabaena* sp. strain PCC 7120 *rpoC1* gene was formed by cloning a 2.1-kb *Xmn*I-*Kpn*I fragment from pT185 into the filled-in *Bam*HI site of pATH3. The *Kpn*I end of the fragment was made blunt with T4 DNA polymerase, and the resulting plasmid with the insert in the proper orientation is pATH301.

Expression of fusion proteins in *E. coli* was carried out as follows. Cultures of MC1061 containing the pATH3 plasmids were grown overnight in M9 medium (27) with 0.5% Casamino Acids (Difco), 50 μg of ampicillin per ml, and 20 μg of tryptophan per ml. The cultures were diluted 1:10 into 5 ml of fresh medium without tryptophan in a 125-ml Erlenmeyer flask and incubated at 30°C with shaking for 1 h. Then 25 μl of a 2-mg/ml solution of 3-β-indoleacrylic acid (Sigma Chemical Co.) in ethanol was added (final concentration, 10 μg/ml), and the cultures were grown for 2 h at 30°C. One milliliter of each culture was pelleted by centrifugation and resuspended in 25 μl of distilled H_2O. An equal amount of 2× final sample buffer was added, and the samples were boiled for 5 min; 20 μl of each sample was loaded on an SDS-polyacrylamide gel, and the gel was blotted to nitrocellulose as described above.

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been assigned GenBank accession number M60831.

RESULTS

Cloning of the *rpoBC1C2* region. RNA polymerase is a highly conserved enzyme, with extensive homology shared between the largest subunits of the enzyme in archaebacteria, eubacteria, and eukaryotes (1, 3, 32, 36, 49). We have used DNA encoding the β and β′ subunits of RNA polymerase in *E. coli* (the *rpoB* and *rpoC* genes, respectively) as a heterologous probe to isolate the analogous genes from *Anabaena* sp. strain PCC 7120. A Southern blot of *Anabaena* sp. strain PCC 7120 chromosomal DNA digested with various restriction enzymes was probed with fragments of the *E. coli rpoB* (Fig. 1A) and *rpoC* genes (not shown). Both probes hybridized to an 11.4-kb *Xba*I fragment and a 6.4-kb *Eco*RI fragment. Rather than using these probes directly to screen two available cosmid libraries for the *Anabaena rpoB* and *rpoC* genes, our plan was to isolate a piece of *Anabaena* DNA to use as a homologous probe, in order to avoid the problem of background hybridization with the *E. coli* DNA present on the cosmid bank filters. Since the results of the Southern blot suggested that the *Anabaena rpoB* and *rpoC* genes were linked, as a matter of convenience we chose to use the *E. coli rpoB* probe to isolate the *Anabaena* 1.3-kb *Hin*dIII-*Eco*RI fragment (Fig. 1). To increase the likelihood of isolating this fragment, we made a size-directed library of *Hin*dIII-*Eco*RI fragments in pBR328 in two steps (see Materials and Methods). Minipreps of DNA from the plasmids were screened by Southern blot analysis using the *E. coli rpoB* probe. pKH100 was one of eight plasmids that hybrid-

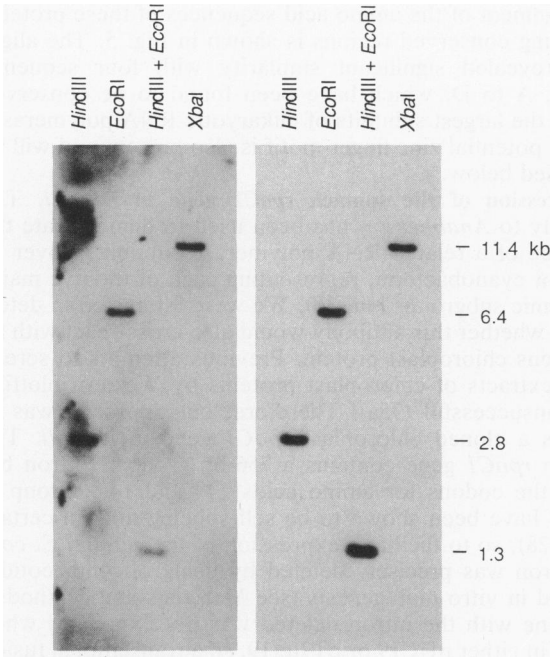

FIG. 1. Southern blot of chromosomal DNA from *Anabaena* sp. strain PCC 7120 digested with various restriction enzymes. (A) Probed with a 2.8-kb *Eco*RI fragment from the *E. coli rpoB* gene; (B) Probed with the 1.3-kb *Eco*RI-*Hin*dIII insert from pKH100.

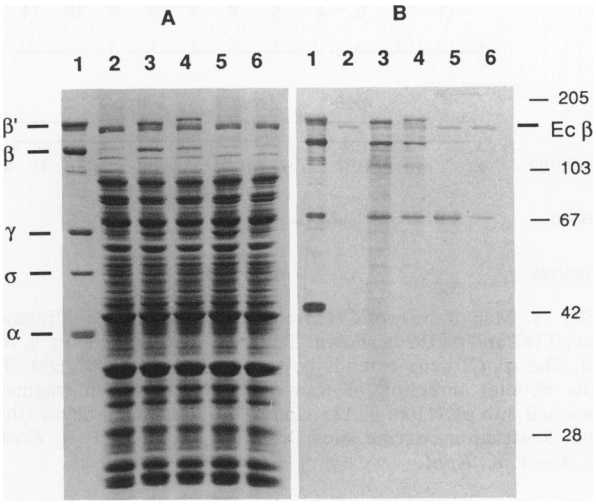

FIG. 2. Expression of *Anabaena* RNA polymerase subunit proteins in *E. coli*. (A) SDS-polyacrylamide gel stained with Coomassie blue; (B) Western blot of an identical gel probed with antiserum to *Anabaena* core RNA polymerase. Sizes of the molecular weight standards are given at right (in kilodaltons). Lanes 1 are purified *Anabaena* RNA polymerase holoenzyme (panel A, 10 μg; panel B, 3 μg). Lanes 2 to 6 are crude extracts of BL21(DE3)pLysS containing the following plasmids: 2, pT7tet18; 3, pT181; 4, pT182; 5, pT195; 6, pT185.

ized to the *rpoB* probe out of 215 tested. The 1.3-kb *Hin*dIII-*Eco*RI insert from pKH100 was used to probe the Southern blot in Fig. 1B, and the same bands hybridized as with the *E. coli rpoB* probe (Fig. 1A). These results indicate that the *rpoB* and *rpoC* genes are present in single copy in *Anabaena* sp. strain PCC 7120.

The 1.3-kb *Hin*dIII-*Eco*RI fragment from pKH100 was used to probe two libraries of *Anabaena* chromosomal DNA in the cosmid vector pWB79 (see Materials and Methods). Seven hybridizing cosmids were identified, and restriction mapping revealed that they were overlapping and contained a region of about 12 kb in common, including the 11.4-kb *Xba*I and 6.4-kb *Eco*RI fragments (see Fig. 1). Southern blot analysis of the cosmid DNAs showed that a probe for the *E. coli rpoC* gene hybridized to the same *Xba*I and *Eco*RI fragments as the *Anabaena rpoB* probe, confirming that the *rpoB* and *rpoC* genes are closely linked in *Anabaena* sp. strain PCC 7120. One of these cosmids, pID9, containing a 35-kb region of *Anabaena* DNA, was used for subsequent experiments.

Expression of RNA polymerase subunits in *E. coli*. To verify that the cloned DNA encoded RNA polymerase subunits, the genes were expressed in *E. coli*. Restriction fragments from the cosmid pID9 were cloned into the expression vector pT7tet18 or pT7tet19 (4), and the resulting plasmids were used to transform *E. coli* BL21(DE3)pLysS (42). This strain contains an inducible T7 RNA polymerase gene under the control of the *lac* UV5 promoter, and the plasmid pLysS encodes T7 lysozyme, which inhibits basal activity of T7 RNA polymerase in the absence of inducer (IPTG). Upon addition of IPTG, the target gene in the plasmid is selectively expressed to a high level. After expression in this system, the *Anabaena* proteins encoded by the plasmids were detected by Western blot analysis using specific antisera to *Anabaena* RNA polymerase subunits as probes (36).

The 11.4-kb *Xba*I fragment from pID9 was subcloned into pT7tet18 to produce pT181 and pT182, which had the insert in opposite orientations. The results of expressing the proteins from these plasmids in *E. coli* are shown in Fig. 2. In the stained gel, proteins with the same relative mobilities as the γ, β, and β′ subunits are clearly visible (Fig. 2A, lanes 3 and 4). The corresponding Western blot shows that the same proteins are bound specifically by antiserum to *Anabaena* core RNA polymerase (Fig. 2B, lanes 3 and 4). The *E. coli* β subunit cross-reacts strongly with the anticore serum (36) and is visible in each lane. Similar amounts of the γ, β, and β′ proteins are produced from pT181 and pT182, and the proteins are also made in the absence of IPTG (data not shown). This finding suggests that expression of the proteins from these plasmids is independent of the T7 promoter and is directed by an *Anabaena* promoter(s) on the *Xba*I fragment. The β′ protein made from pT181 is slightly smaller than the native β′, and DNA sequencing has shown that the *rpoC2* gene encoding β′ extends 215 bp past the end of the *Xba*I fragment (3a). The β′ protein produced by pT182 is larger because of translational read-through into the vector sequences.

To localize the genes encoding the various subunits, smaller fragments were subcloned from the 11.4-kb *Xba*I fragment. By screening Western blots of proteins expressed by the subclones with antibodies to the γ subunit and the β plus β′ subunits of RNA polymerase, the organization of the genes was deduced (Fig. 3). For example, the 7.4-kb *Xba*I-*Kpn*I fragment produced the β and γ proteins, while the 6.4-kb *Eco*RI fragment made γ and a truncated β′ peptide (data not shown). A 4.6-kb *Hin*dIII fragment internal to the *Eco*RI fragment encoded γ and a smaller truncated β′ (data not shown). Finally, a 2.8-kb *Hin*dIII-*Kpn*I fragment, cloned

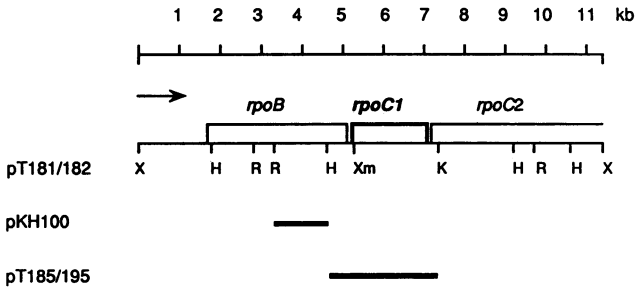

FIG. 3. Map of the *rpoBC1C2* region. The 11.4-kb *Xba*I fragment from pT181 and pT182 is shown. The *rpoC1* gene, encoding γ, is in bold. The *rpoC2* gene extends beyond the *Xba*I site at right. The arrow denotes direction of transcription. Restriction fragments subcloned into pKH100, pT185, and pT195 are shown below (thick lines). Restriction enzyme sites: X, *Xba*I; H, *Hin*dIII; R, *Eco*RI; Xm, *Xmn*I; K, *Kpn*I.

in opposite orientations in pT7tet18 and pT7tet19 to make pT185 and pT195, produced only γ (Fig. 2, lanes 5 and 6). These experiments established that the genes for the β, γ, and β′ subunits are arranged in the order *rpoB*, *rpoC1*, *rpoC2*. The gene for the β subunit is called *rpoB* because of homology with the *E. coli rpoB* gene. The genes for γ and β′ of *Anabaena* sp. are called *rpoC1* and *rpoC2*, respectively, because it appears that the functions of the *E. coli* β′ subunit are split in *Anabaena* sp. between the γ and β′ subunits (see Discussion). In Fig. 2, note that synthesis of the γ subunit is directed by both pT185 and pT195, but γ is more highly expressed from pT195 (lanes 5). In general, the fragments subcloned from the 11.4-kb *Xba*I fragment directed the synthesis of more protein in one orientation than the other, which established that the direction of transcription is from left to right on the map in Fig. 3.

DNA sequence of the *rpoC1* gene. The DNA sequence of the 2.8-kb *Hin*dIII-*Kpn*I fragment was determined and is shown in Fig. 4. Starting at the *Hin*dIII site, there is part of an open reading frame encoding a 165-amino-acid peptide 70% homologous to the carboxy-terminal end of *E. coli* β. This is the end of the *rpoB* gene. After a 133-bp intergenic space, another open reading frame of 1,875 bp follows, encoding a 625-amino-acid protein with a molecular weight of 70,562. The sequence of the N-terminal five amino acids of purified γ protein was determined at the Harvard University Microchemistry Facility, and the results confirmed the start of this open reading frame as the start of γ. A third open reading frame begins 100 bp after the end of *rpoC1*. It encodes a 63-amino-acid peptide which is about 28% identical to a region beginning in the C-terminal half of *E. coli* β′. This is most likely the start of the *rpoC2* gene. The *rpoC2* gene is preceded by a DNA sequence that has a good match to a Shine-Dalgarno ribosome binding site. The match is not as good with the corresponding sequence preceding the start of *rpoC1*.

Homology of the γ subunit with RNA polymerase subunits from other organisms. The γ subunit of *Anabaena* RNA polymerase corresponds to the amino-terminal 590 amino acids of the β′ subunit of *E. coli* RNA polymerase, with about 60% identity at the amino acid level. There is approximately 96% identity between the γ subunits of *Anabaena* sp. strain PCC 7120 and the closely related cyanobacterium *N. commune* UTEX 584. Plant chloroplasts contain an analogous RNA polymerase subunit (β′) (19, 20, 38). The spinach chloroplast β′ is about 44% identical to *Anabaena* γ. An alignment of the amino acid sequences of these proteins indicating conserved regions is shown in Fig. 5. The alignment revealed significant similarity with four sequence blocks, A to D, which have been found to be conserved among the largest subunits of eukaryotic RNA polymerases (21). A potential zinc finger motif is also present and will be discussed below.

Expression of the spinach *rpoC1* gene in *E. coli*. The antibody to *Anabaena* γ has been used to demonstrate the presence of a related RNA polymerase subunit in over 30 different cyanobacteria, representing each of the five major taxonomic subgroups (4a, 36). We were interested in determining whether this antibody would also cross-react with an analogous chloroplast protein. Previous attempts to screen crude extracts of chloroplast proteins by Western blotting were unsuccessful (35a). Therefore, our approach was to express a cloned chloroplast *rpoC1* gene in *E. coli*. The spinach *rpoC1* gene contains a 756-bp group II intron between the codons for amino acids 144 and 145. Group II introns have been shown to be self-splicing only in certain cases (28), so to facilitate expression of the gene in *E. coli*, the intron was precisely deleted by using oligonucleotide-directed in vitro mutagenesis (see Materials and Methods). The gene with the intron deleted was not expressed when cloned in either pUC19 or pT7tet19, so a translational fusion was made with the truncated *E. coli trpE* gene in the vector pATH3 (23). This vector was chosen because it allows high-level inducible expression of fused proteins under control of the strong *trp* promoter and translation initiation region. The fusion removed only the first two amino acids of the chloroplast β′ subunit, and when joined with the *trpE* fragment of 37 kDa, a fusion product of about 115 kDa was expected (pATH302). As a control for expression, a fusion was also made to the *Anabaena rpoC1* gene. The first 21 amino acids of γ were removed, and the expected size of the fusion product was about 105 kDa (pATH301). A fusion was made with the undeleted chloroplast *rpoC1* gene as well, to determine whether self-splicing of the intron occurs in *E. coli*(pATH303).

The results of protein expression from the pATH plasmids are shown in Fig. 6. The stained gel (Fig. 6A) shows that from each plasmid a major polypeptide product was synthesized. The vector produced the 37-kDa truncated *trpE* protein (lane 1), while pATH301 and pATH302 produced fusion proteins of the expected sizes (lanes 2 and 3). The strong band of about 53 kDa made from pATH303 in lane 4 results from translation of only the first exon of the intron-containing chloroplast *rpoC1* gene. The plasmid with the intron deleted (pATH302) consistently made less fusion product than the other constructions. The reason for the lowered expression is unclear, but it may be due to problems with codon usage or protein degradation.

The corresponding Western blot (Fig. 6B) shows that the fusion protein made by the intron-deleted plasmid pATH302 does cross-react with the anti-γ serum (lane 3). In lane 4, a weak reaction of the antibody with the product of the first exon of the chloroplast gene is apparent. A larger protein of the size predicted for the fusion of full-length β′ is not present in this lane, indicating that self-splicing of the intron probably does not occur in *E. coli* or does so inefficiently.

DISCUSSION

We have isolated the genes for the three largest subunits of *Anabaena* sp. strain PCC 7120 RNA polymerase. The genes are linked in the order *rpoB*, *rpoC1*, *rpoC2* and encode the β,

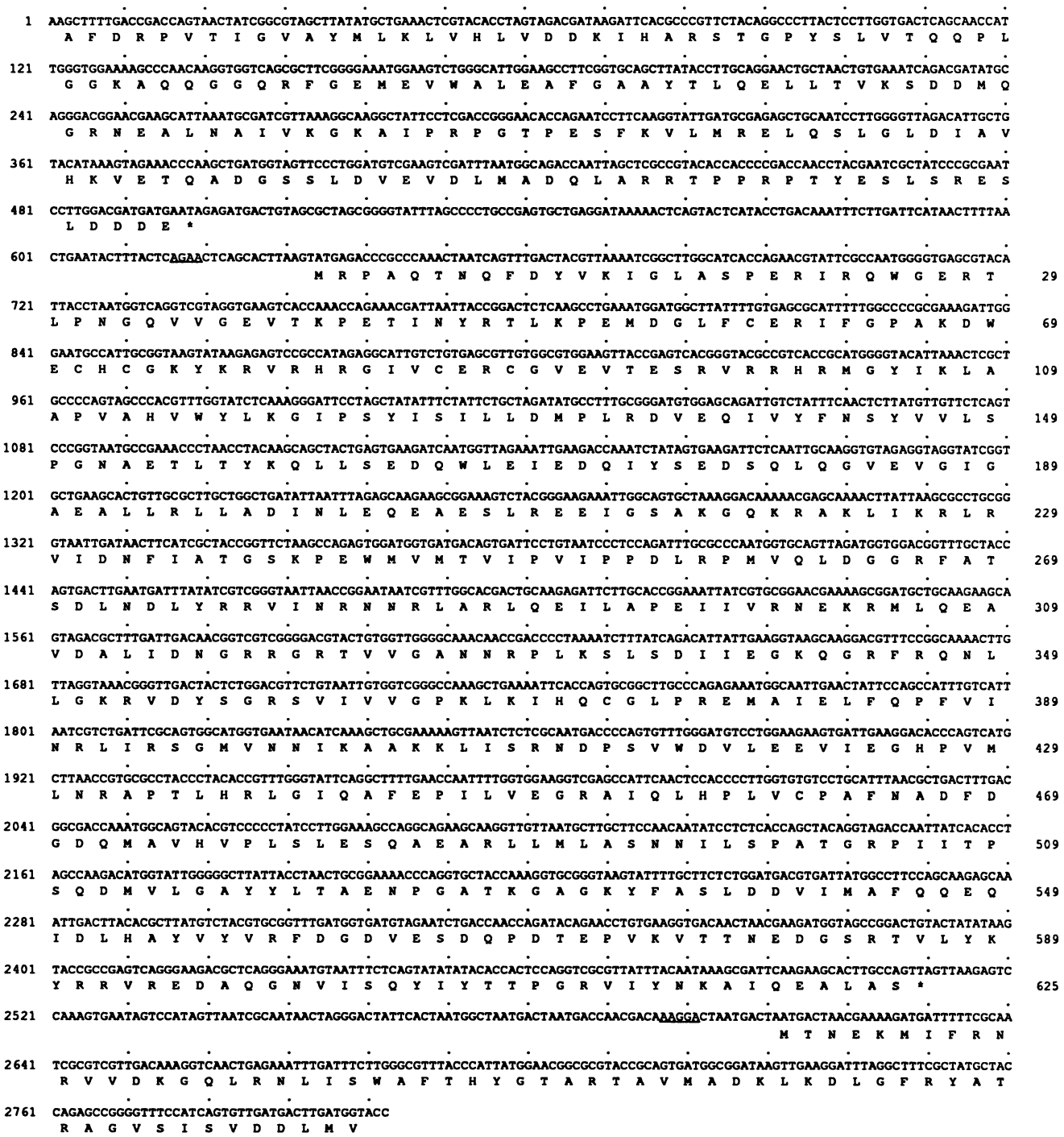

FIG. 4. Nucleotide and derived amino acid sequences of the *Anabaena* sp. strain PCC 7120 *rpoC1* gene. Nucleotide numbers are at left; amino acid numbers of the γ protein are at right. The end of the *rpoB* reading frame extends from nucleotides 1 to 497. The *rpoC2* reading frame begins at nucleotide 2612. Potential ribosome binding sites are underlined.

γ, and β′ subunits, respectively. The predicted molecular sizes of γ (70 kDa) and β (124 kDa) agree with previous estimates (37), but the β′ subunit is predicted to be slightly smaller (159 kDa) than the 171 kDa reported. The genes are colinear with the *rpoB* and *rpoC* genes of *E. coli*, but homology to the β′ subunit of *E. coli* is divided between two genes, *rpoC1* and *rpoC2*. The same split in the *rpoC* gene is present in the cyanobacterium *N. commune* UTEX 584 (47) and in plant chloroplasts (19, 20, 38). While the organization of the cyanobacterial and chloroplast *rpoBC1C2* genes resembles the *rpoBC* operon of *E. coli*, the *Anabaena* and chloroplast genes do not lie distal to any ribosomal protein genes (3a, 19, 20). The regulation of expression of the *Anabaena rpoBC1C2* genes is not yet clear, but it must be

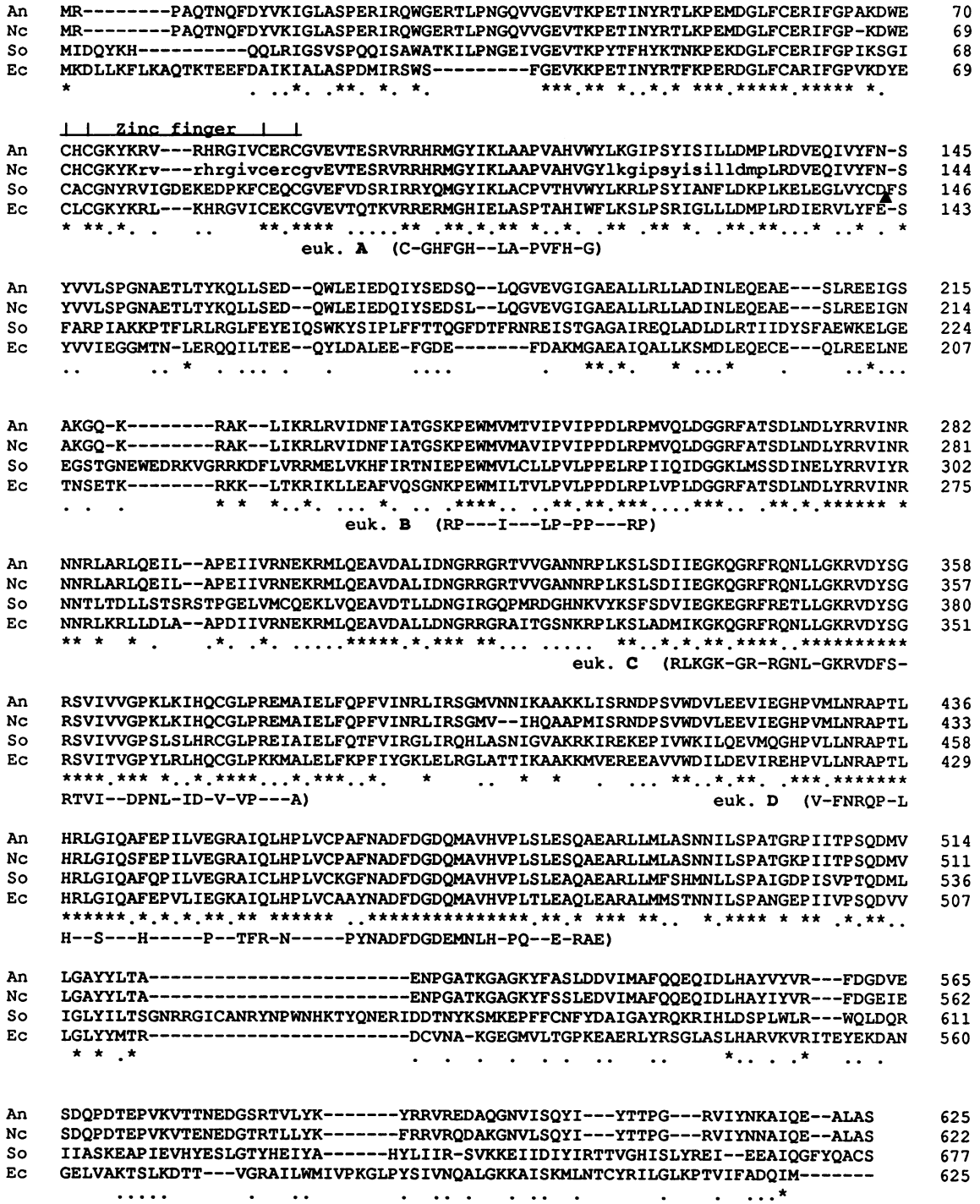

FIG. 5. Alignment of the predicted amino acid sequences of several RNA polymerase γ and β′ subunits. Included are the γ proteins from *Anabaena* sp. strain PCC 7120 (An) and *N. commune* UTEX 584 (47) (Nc) and the β′ proteins from spinach chloroplast (19) (So) and *E. coli* (31) (Ec). The comparison covers the first 625 amino acids of the *E. coli* β′ subunit. Amino acid numbers of the proteins are at right. Residues in lowercase letters (Nc, amino acids 77 to 90 and 118 to 132) differ from the published sequence; these residues were in an alternate reading frame, indicative of possible sequence errors. An asterisk denotes amino acid identity, while a dot represents a conservative substitution. The position of the intron in the spinach sequence is marked by a filled triangle. The eukaryotic homology blocks A to D (21) are indicated, with dashes representing gaps in the eukaryotic consensus. A potential zinc finger motif is overlined, with vertical lines denoting conserved cysteines. Alignments were made by using the Clustal program for generating multiple sequence alignments (17).

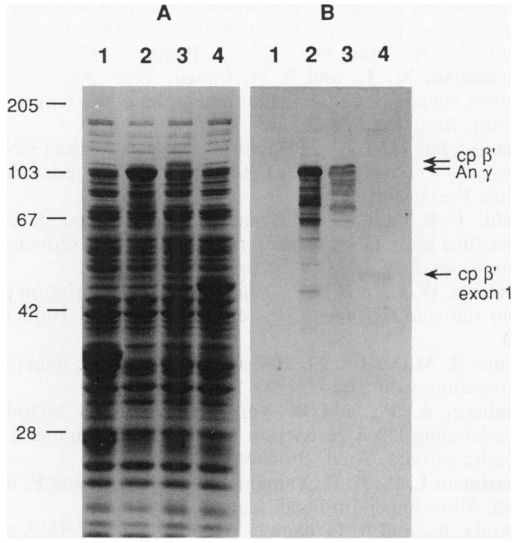

FIG. 6. Expression of the spinach chloroplast *rpoC1* gene in *E. coli* as a fusion protein. (A) SDS-polyacrylamide gel stained with Coomassie blue. Sizes of molecular weight standards are given at left (in kilodaltons). (B) Western blot of an identical gel probed with antiserum to the *Anabaena* γ subunit. Samples were crude extracts of MC1061 containing the following plasmids: lanes 1, pATH3; lanes 2, pATH301; lanes 3, pATH302; and lanes 4, pATH303. Positions of the fusion proteins are at the right (An γ is *Anabaena* γ; cp β' is chloroplast β' protein).

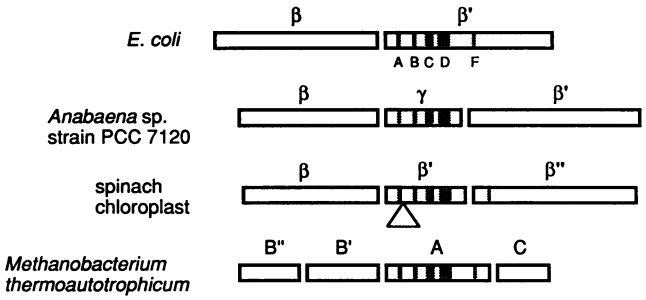

FIG. 7. Arrangement of the genes encoding the largest subunits of RNA polymerase in eubacteria, cyanobacteria, plant chloroplasts, and archaebacteria. Shown are the genes from *E. coli* (31), *Anabaena* sp. strain PCC 7120 (this study), spinach chloroplast (19), and *M. thermoautotrophicum* (3). Genes are aligned with respect to the start of the β'-like sequences. Protein products are indicated above each gene. Triangle below spinach chloroplast β' indicates position of the intron. Shaded boxes represent conserved homology blocks A to F (21).

different from that in *E. coli*, in which the *rpoBC* genes are transcribed from an upstream ribosomal protein gene promoter (9).

We have shown that the *rpoC1* gene of spinach chloroplasts can be expressed in *E. coli* and that it encodes a protein immunologically related to *Anabaena* γ. The gene is expressed in *E. coli* only if the 756-bp group II intron is deleted. This finding suggests that the intron either is not self-splicing or is inefficiently spliced in *E. coli*. The *rpoA*, *rpoB*, and *rpoC2* genes of various plant chloroplasts have previously been shown to produce protein products (18, 33, 34), but this is the first demonstration of a protein expressed from a chloroplast *rpoC1* gene. The failure of previous attempts to detect a protein related to γ in crude extracts of chloroplasts is not due to a lack of cross-reactivity with the anti-γ antibody, but rather to a low abundance of the protein in the extracts. Since anti-γ does indeed cross-react with the product of the spinach *rpoC1* gene, it may prove to be a useful tool for the study of chloroplast RNA polymerase.

The cyanobacterial γ and chloroplast β' subunits are homologous to the amino-terminal half of the *E. coli* β' subunit (this study; 19, 20, 47). Four regions of similarity correspond to eukaryotic consensus homology blocks A to D, which are conserved among the largest subunits of eukaryotic RNA polymerases (21) (Fig. 5). The most striking regions of sequence similarity surround the GKRVD motif in block C (γ amino acids 351 to 355) and the NADFDGD motif in block D (γ amino acids 465 to 471). The latter represents the longest conserved region which is universally present in all of the largest subunits of eukaryotic RNA polymerases (22). The function of these regions is unclear, but the high degree of conservation implies that they identify domains of functional importance. It has been suggested that region C has segments which make up a helix-turn-helix motif similar to a region in *E. coli* DNA polymerase I thought to be important for DNA binding (1). However, a computer search for a helix-turn-helix motif in the γ protein by using the method of Dodd and Egan (8) did not predict such a structure in this region.

Another conserved region that may be functionally important is a cysteine-rich region located just upstream of block A. As noted by Hudson et al. (19) for spinach chloroplast β', there are four cysteine residues within 22 amino acids that could possibly form a zinc-binding pocket. RNA polymerase is a zinc-containing enzyme, and the β' subunit of *E. coli* RNA polymerase contains at least one of the two Zn(II)-binding sites (29, 46). The *E. coli* β' subunit is also thought to be involved in binding the DNA template (13). Therefore, this motif could be analogous to a C_4-type zinc finger structure, implicated in DNA binding by eukaryotic transcription factors (10, 41). The prokaryotic motif C-X-C-X_{12}-C-X_2-C does not perfectly match the eukaryotic C_4-type consensus C-X_2-C-X_{13}-C-X_2-C. However, it has been proposed that sequence differences among zinc fingers may reflect variations on a common structural theme, allowing a diversity of functions and an alteration or relaxation of DNA sequence specificity (12, 20). Alternatively, it is possible that the putative zinc-binding region functions catalytically in transcription or plays a structural role in maintaining the proper subunit arrangement of the enzyme (21, 46).

In the archaebacteria, homology to *E. coli* β' is also split between two subunits, A and C. In some cases, the homology to β is split as well. The genes for these subunits are linked and colinear with the corresponding eubacterial and chloroplast genes (Fig. 7) (3, 32, 49). Serological studies have shown that antiserum to *Anabaena* γ cross-reacts with the A subunit of *Sulfolobus acidocaldarius* (36). A comparison of the amino acid sequences of γ and the A subunit of *Methanobacterium thermoautotrophicum* (3) revealed only about 18% identity (data not shown), but homology with the major conserved regions A to D, as well as a zinc-binding motif, was clearly present. However, the γ and A subunits differ in the extent of the region of sequence similarity spanning *E. coli* β'. The γ subunit corresponds to the first one-half of *E. coli* β', while the A subunit is homologous to approximately the first two-thirds of β' (32), including conserved homology block F (21) (Fig. 7). Block F is not present in γ or chloroplast β' but is found in the chloroplast β"

protein instead. Most likely it will also be found in the *Anabaena* β′ subunit when the sequence of the *rpoC2* gene is completed. Thus, the sequence homologous to *E. coli* β′ is divided at a different location in archaebacteria than in cyanobacteria and chloroplasts. What can be said about the structure of the ancestral RNA polymerase β′ subunit? It might have consisted of a single polypeptide whose gene split in different places during the evolution of the cyanobacterial and archaebacterial lineages. Alternatively, either of the split genes might have fused to give rise to the ancestor of the contemporary *rpoC* gene of other eubacteria.

The gene organization and subunit structure of RNA polymerase may be useful in establishing phylogenetic relationships between organisms. Comparisons of rRNA sequences have been used for this purpose (14, 45), but RNA polymerase may provide a good alternative because it is ubiquitous, highly conserved, and a complex macromolecule (32). Archaebacterial RNA polymerases have a more complex component pattern than their eubacterial counterparts, comprising approximately 10 subunits. In structure, they more closely resemble eukaryotic nuclear RNA polymerases (32). Studies using the β′-like subunits of RNA polymerases to determine phylogenetic relationships among archaebacteria, eubacteria, and eukaryotes have found that the archaebacteria are a coherent group and are indeed closely related to the eukaryotic nuclear RNA polymerase II and III lineages (32, 49). In addition, eubacteria and chloroplasts were found to form another distinct group.

The organization of RNA polymerase genes in chloroplasts clearly reflects a prokaryotic operon structure (19). Likewise, the striking similarity in the RNA polymerase gene products and gene organizations between cyanobacteria and chloroplasts provides further evidence for a common evolutionary pathway among cyanobacteria and chloroplasts and a divergence of this group from the rest of the eubacterial kingdom. The evolution of the β′ subunit of RNA polymerase supports the endosymbiont hypothesis of cyanobacteria as progenitors of plant chloroplasts.

ACKNOWLEDGMENTS

We thank William Buikema for use of his cosmid libraries, Paul Gardner for synthesis of oligonucleotides, and Bianca Brahamsha for preparation of proteins for N-terminal sequencing. We are grateful to William Buikema, Bianca Brahamsha, and Brian Palenik for critical reading of the manuscript.

This work was supported by a postdoctoral fellowship (grant PF-3235) from the American Cancer Society to K.J.B. and by Public Health Service grant GM21823 from the National Institutes of Health.

REFERENCES

1. **Allison, L. A., M. Moyle, M. Shales, and C. J. Ingles.** 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell **42**:599–610.
2. **Barry, G., C. L. Squires, and C. Squires.** 1979. Control features within the *rplJL-rpoBC* transcription unit of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **76**:4922–4926.
3. **Berghöfer, B., L. Kröckel, C. Körtner, M. Truss, J. Schallenberg, and A. Klein.** 1988. Relatedness of archaebacterial RNA polymerase core subunits to their eubacterial and eukaryotic equivalents. Nucleic Acids Res. **16**:8113–8128.

3a. **Bergsland, K. J.** Unpublished data.

4. **Borthakur, D., M. Basche, W. J. Buikema, P. B. Borthakur, and R. Haselkorn.** 1990. Expression, nucleotide sequence and mutational analysis of two open reading frames in the *nif* gene region of *Anabaena* sp. strain PCC 7120. Mol. Gen. Genet. **221**:227–234.

4a. **Brahamsha, B.** Unpublished data.

5. **Buikema, W. J., and R. Haselkorn.** 1991. Isolation and complementation of nitrogen fixation mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **173**:1879–1885.
6. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
7. **Chamberlin, M. J.** 1982. Bacterial DNA-dependent RNA polymerases, p. 61–86. *In* P. D. Boyer (ed.), The enzymes. Academic Press, New York.
8. **Dodd, I. B., and J. B. Egan.** 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. **18**:5019–5026.
9. **Downing, W. L., and P. P. Dennis.** 1987. Transcription products from the *rplKAJL-rpoBC* gene cluster. J. Mol. Biol. **194**:609–620.
10. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc fingers: gilt by association. Cell **52**:1–3.
11. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
12. **Freedman, L. P., K. R. Yamamoto, B. F. Luisi, and P. B. Sigler.** 1988. More fingers in hand. Cell **54**:444.
13. **Fukuda, R., and A. Ishihama.** 1974. Subunits of RNA polymerase in function and structure. V. Maturation in vitro of core enzyme from *E. coli*. J. Mol. Biol. **87**:523–540.
14. **Giovannoni, S. J., S. Turner, G. J. Olsen, S. Barns, D. J. Lane, and N. R. Pace.** 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. **170**:3584–3592.
15. **Heinrich, P., B. T. Nixon, and D. Voytas.** 1988. Construction of nested deletions for DNA sequencing, p. 7.3.1–7.3.21. *In* F. A. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. D. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing and Wiley-Interscience, New York.
16. **Helmann, J. D., and M. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57**: 839–872.
17. **Higgins, D. G., and P. M. Sharp.** 1989. Clustal: a package for performing multiple sequence alignments on a microcomputer. Gene **73**:237–244.
18. **Hu, J., and L. Bogorad.** 1990. Maize chloroplast RNA polymerase: the 180-, 120-, and 38-kilodalton polypeptides are encoded in chloroplast genes. Proc. Natl. Acad. Sci. USA **87**:1531–1535.
19. **Hudson, G. S., T. A. Holton, P. R. Whitfeld, and W. Bottomley.** 1988. Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. J. Mol. Biol. **200**:639–654.
20. **Igloi, G. L., A. Meinke, I. Döry, and H. Kössel.** 1990. Nucleotide sequence of the maize chloroplast *rpoB/C1/C2* operon: comparison between the derived protein primary structures from various organisms with respect to functional domains. Mol. Gen. Genet. **221**:379–394.
21. **Jokerst, R. S., J. R. Weeks, W. A. Zehring, and A. L. Greenleaf.** 1989. Analysis of the gene encoding the largest subunit of RNA polymerase II in *Drosophila*. Mol. Gen. Genet. **215**:266–275.
22. **Köck, J., R. Evers, and A. W. C. A. Cornelissen.** 1988. Structure and sequence of the gene for the largest subunit of trypanosomal RNA polymerase III. Nucleic Acids Res. **16**:8753–8772.
23. **Koerner, T. J., J. E. Hill, A. M. Myers, and A. Tzagoloff.** 1990. High expression vectors with multiple cloning sites for construction of *trpE* fusion genes: pATH vectors. Methods Enzymol. **194**:477–490.
24. **Kraft, R., J. J. Tardiff, K. S. Krauter, and L. A. Leinwand.** 1988. Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase. BioTechniques **6**:544–547.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
26. **Lerbs, S., E. Bräutigam, and B. Parthier.** 1985. Polypeptides of DNA-dependent RNA polymerase of spinach chloroplasts: characterization by antibody-linked polymerase assay and determination of sites of synthesis. EMBO J. **4**:1661–1666.
27. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory,

Cold Spring Harbor, N.Y.
28. **Michel, F., K. Umesono, and H. Ozeki.** 1989. Comparative and functional anatomy of group II catalytic introns—a review. Gene **82**:5–30.
29. **Miller, J. A., G. F. Serio, R. A. Howard, J. L. Bear, J. E. Evans, and A. P. Kimball.** 1979. Subunit localizations of zinc(II) in DNA-dependent RNA polymerase from *E. coli* B. Biochim. Biophys. Acta **579**:291–297.
30. **Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, S. Sano, K. Umesono, Y. Shiki, M. Takeuchi, Z. Chang, S. Aota, H. Inokuchi, and H. Ozeki.** 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature (London) **322**:572–574.
31. **Ovchinnikov, Y. A., G. S. Monastyrskaya, V. V. Gubanov, S. O. Guryev, I. S. Salomatina, T. M. Shuvaeva, V. M. Lipkin, and E. D. Sverdlov.** 1982. The primary structure of *E. coli* RNA polymerase. Nucleotide sequence of the *rpoC* gene and amino acid sequence of the β′ subunit. Nucleic Acids Res. **10**:4035–4044.
32. **Pühler, G., H. Leffers, F. Gropp, P. Palm, H.-P. Klenk, F. Lottspeich, R. Garrett, and W. Zillig.** 1989. Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc. Natl. Acad. Sci. USA **86**:4569–4573.
33. **Purton, S., and J. C. Gray.** 1989. The plastid *rpoA* gene encoding a protein homologous to the bacterial RNA polymerase alpha subunit is expressed in pea chloroplasts. Mol. Gen. Genet. **217**:77–84.
34. **Ruf, M., and H. Kössel.** 1988. Structure and expression of the gene coding for the α-subunit of DNA-dependent RNA polymerase from the chloroplast genome of *Zea mays*. Nucleic Acids Res. **16**:5741–5754.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
35a. **Schneider, G.** Unpublished data.
36. **Schneider, G. J., and R. Haselkorn.** 1988. RNA polymerase subunit homology among cyanobacteria, other eubacteria, and archaebacteria. J. Bacteriol. **170**:4136–4140.
37. **Schneider, G. J., N. E. Tumer, C. Richaud, G. Borbely, and R. Haselkorn.** 1987. Purification and characterization of RNA polymerase from the cyanobacterium *Anabaena* 7120. J. Biol. Chem. **262**:14633–14639.
38. **Shimada, H., M. Fukuta, M. Ishikawa, and M. Sugiura.** 1990. Rice chloroplast RNA polymerase genes: the absence of an intron in *rpoC1* and the presence of an extra sequence in *rpoC2*. Mol. Gen. Genet. **221**:395–402.
39. **Shinozaki, K., M. Ohme, M. Tanaka, T. Wakasugi, N. Hayashida, T. Matsubayashi, N. Zaita, J. Chunwongse, J. Obokata, K. Yamaguchi-Shinozaki, C. Ohto, K. Torazawa, B. Y. Meng, M. Sugita, H. Deno, T. Kamogashira, K. Yamada, J. Kusuda, F. Takaiwa, A. Kato, N. Tohdoh, H. Shimada, and M. Sugiura.** 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO. J. **5**:2043–2049.
40. **Sijben-Müller, G., R. B. Hallick, J. Alt, P. Westhoff, and R. G. Herrman.** 1986. Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. Nucleic Acids Res. **14**:1029–1044.
41. **Struhl, K.** 1989. Helix-turn-helix, zinc-finger and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. Trends Biochem. Sci. **14**:137–140.
42. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
43. **Tortorello, M. L., and G. M. Dunny.** 1985. Identification of multiple cell surface antigens associated with the sex pheromone response of *Streptococcus faecalis*. J. Bacteriol. **162**:131–137.
44. **von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen.** 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. **53**:389–446.
45. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**:221–271.
46. **Wu, C.-W., F. Y.-H. Wu, and D. C. Speckhard.** 1977. Subunit location of the intrinsic divalent metal ions in RNA polymerase from *E. coli*. Biochemistry **16**:5449–5454.
47. **Xie, W. Q., K. Jager, and M. Potts.** 1989. Cyanobacterial RNA polymerase genes *rpoC1* and *rpoC2* correspond to *rpoC* of *Escherichia coli*. J. Bacteriol. **171**:1967–1973.
48. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
49. **Zillig, W., H.-P. Klenk, P. Palm, G. Pühler, F. Gropp, R. A. Garrett, and H. Leffers.** 1989. The phylogenetic relations of DNA-dependent RNA polymerases of archaebacteria, eukaryotes, and eubacteria. Can. J. Microbiol. **35**:73–80.